



TITLE:

適応する記号力学系と知能について: 自律ロボットによる行動学習の実験(ポスター発表, 基研長期研究会「複雑系」, 研究会報告)

AUTHOR(S):

谷, 淳

CITATION:

谷, 淳. 適応する記号力学系と知能について: 自律ロボットによる行動学習の実験(ポスター発表, 基研長期研究会「複雑系」, 研究会報告). 物性研究 1995, 63(6): 850-855

ISSUE DATE:

1995-03-20

URL:

<http://hdl.handle.net/2433/95493>

RIGHT:

適応する記号力学系と知能について: 自律ロボットによる行動学習の実験

谷 淳

ソニーコンピュータサイエンス研究所

Self-Organization of Symbols in Behavior-Based Robot Learning

Jun Tani, tani@csl.sony.co.jp, Sony CSL

Abstract

The paper describes how a dynamical system's approach can support intelligent computations through experiments of an autonomous mobile robot. The robot interacts with the physical world by means of its motor outputs and local sensory inputs, through which it learns an internal model of the world as embedded on its internal neural dynamics. It turns out that the grammatical representation that describes the rule set of the world is acquired, which enables the robot to conduct a deliberative computation such as planning. It is concluded that symbolic representations and their manipulations are self-organized through the behavioral experiences.

1 概要

本文では、知性とはカオスを含む複雑系の為す技であるといった解釈を、現実のロボットによる実験をふまえながら示す。ロボットは、物理的世界に対してモータコマンドを発生し、それに対するセンサー情報を受け、その結果としてまたモータコマンドを出すといった、一つのダイナミカルループを形成する。したがって、ロボットの挙動はこの内部と環境の力学系の結合により決定されることになる。この時、ロボットの知性のレベルは、この全体の力学系が複雑さのどの階層に位置されるかによって、分類可能と考えられる。

さて、ロボットが行動を通して何らかの知

識を学習していく過程は、その内部力学系を変えていく過程に他ならない。一般的に、学習によって得られた知識の表現及びその活用には二通りあると思われる。一つは、「スキル学習」といわれるもので、それはルックアップテーブルを用意し、経験して得られた内部状態と取るべき行動のマッピングを、そこに埋めていくことに等しい。ロボットはこのテーブルから、各時点での内部状態に対応する行動を決定していき、達成すべきゴール状態へと進んでいく。この方法では、達成すべきゴール状態を任意に変えることはできない。つまり、一つの目的のスキルを学習した場合、それを他の目的に連用することは基本的にはできない。もう一つの方法は「モデル学習」、つまり行動に基づくところの、ロボットと環境の因果関係モデルを獲得していくといったものである。モデルの記述は、状態 x × 行動 y → 状態 z といった文法規則集合的な構造を持ちうり、ロボットはそれらの組み合わせ操作をすることにより、任意のゴール状態への行動プランを立案することができる。筆者は、知性とは、モデルを用い多様な問題にアプローチできるメカニズムを指し示すと考え、本文ではモデルの獲得、モデルに基づくプランの立案といった計算過程を、いかに力学系の枠組の上で実現していくか説明する。そして、知性の本質ともいえる、Deliberative な計算過程はカオス力学系の時間発展に対応することを示す。

詳細な説明に入る前に、最後に力学系に基づく筆者のアプローチと伝統的 AI に基づくその違いを説明する。AI 的手法は、基本的には、記号的知識の表現とその操作の上に成

り立っていると言えよう。カオス力学系においても、記号力学系の理論が示す通りに、同様の記号操作の過程を考えることができる。但し、AI 的手法では記号表現はプログラマーが与えるが、筆者の主張する方法では、記号自身が内部力学系の環境との物理的干渉の過程に自己組織化されるという大きな違いがある。筆者のロボットは、それ自身の行動に基づくより直接的な記号的表現および操作を自ら作り出すといえる。いかに世界を、操作可能なアブストラクたなかたちで、しかも現実との致命的なギャップを作り出さずに、表現していくかは長年の問題であるが、筆者の主張する力学系に基づく方法論は、それに対して、ひとつの方向を与えるものである。

2 移動ロボット

理論的な定式化を行なう前に、移動ロボットの説明をする。筆者らは、移動ロボットヤマビコにレーザ型周囲距離センサーを、取り付け(図1参照)、ほぼリアルタイムで距離画像を得られるようにした。

以下に、ナビゲーションのアーキテクチャーについて説明する。ロボットのナビゲーションは、コントロールレベルとナビゲーションレベルの二階層で行なわれる。コントロールレベルでは、距離画像を用いたポテンシャル法に基づく、衝突回避制御が行なわれる。ロボットは基本的には、距離画像ポテンシャル曲線のローカルピークをトラックしながら、スムーズに前進していく。図2に、ロボットが作業空間を移動していく過程に、センサー情報がどのように変化していくか、時系列的に示す。ロボットがあるローカルピークをトラックして移動していく過程に、別のローカルピーク、つまり新たな別の進める方向が発生しうる。この時、ナビゲーションレベルは、現在のローカルピークを追跡し続けるか、新たな分岐方向に進むか判断する。適切な分岐判断を繰り返していくことにより、ロボットは目的地に到達する。本手法において、ナビゲーションの問題は、分起点での方向の選択の問題に集約されている。本章の最後に、ノイズーな実世界での実験において、各分起点は一つの点でなく、確率的分散を持つことを記しておく。

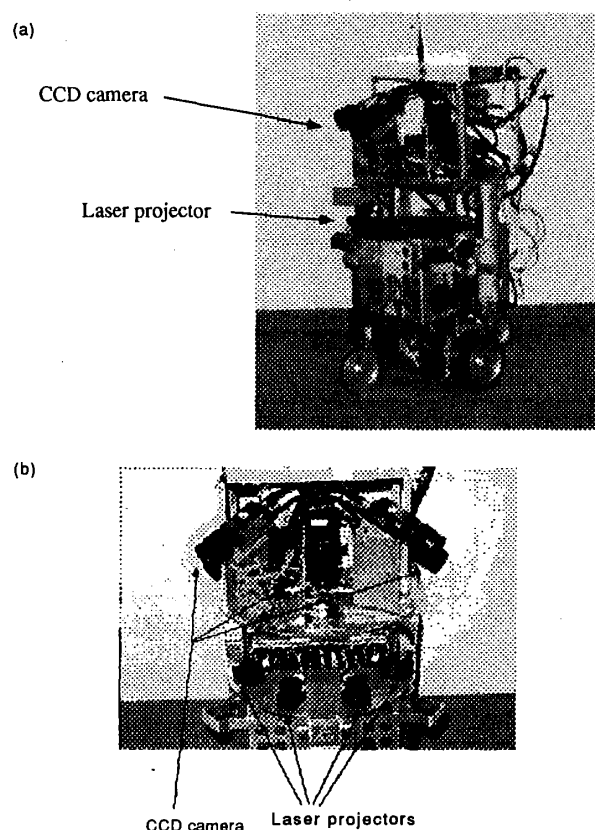


Figure 1: (a) shows the mobile robot YAM-ABICO, (b) shows its laser range sensor consisting of three CCD camera and laser line projectors.

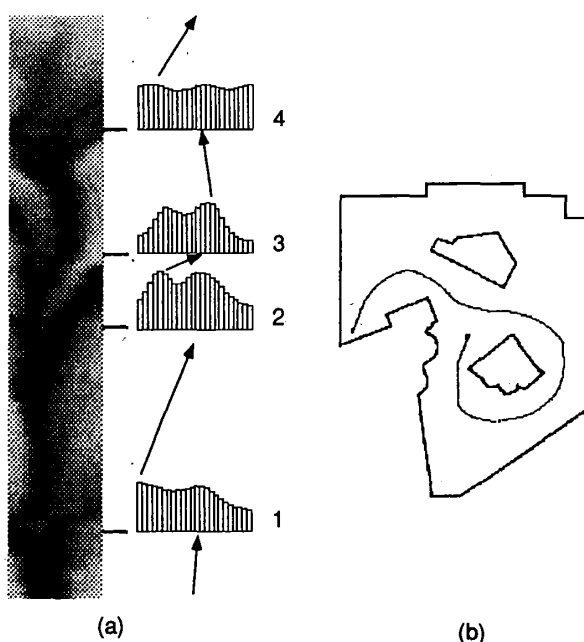


Figure 2: An example travel and its sensory flow.

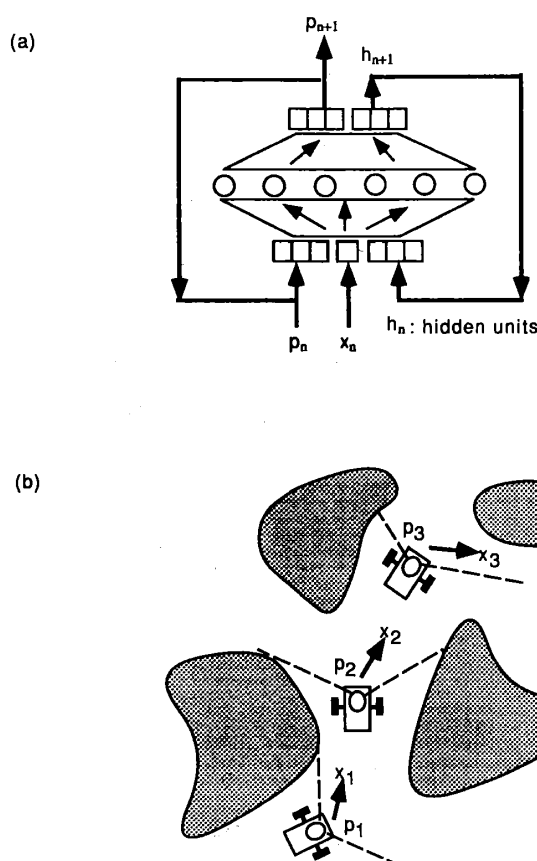


Figure 3: (a) is an RNN architecture for forward modeling, (b) is schematics of how a robot obtains sensori-motor sequences.

3 定式化

以下に、力学系の枠組を用いた、定式化を示す。まず移動ロボットのセンサー情報に基づく、モデルの獲得について川人らの提唱するフォワードモデルの枠組から説明していく。そして、学習して得られたフォワードモデルから、どのように目的のモータプランを生成するか、その時なぜカオス力学系が必要となるか説明する。

3.1 フォワードモデル

学習の目的は、適当なモータプラン（この場合、分岐方向の選択の時系列）について、それがどのようなセンサー入力時系列が発生し得るか予測するようなフォワードモデルを、作り上げることである。そのようなフォワードモデルは、時間ステップ n における内部状態を s_n 、モータコマンドを x_n 、センサー入力を p_n とした場

合、次の時間発展方程式にて表現できる。

$$s_{n+1} = f_s(s_n, x_n) \quad (1)$$

$$p_{n+1} = f_p(s_n, x_n) \quad (2)$$

この時間発展方程式を、図 3 に示すようなリカレントニューラルネット (RNN) を用いて、表現する。このニューラルネット上の定式化において、 x_n はセンサー入力ユニットと隠れ状態ユニットの和により表される。この隠れ状態ユニットがあることにより、文脈依存的ナビゲーションが可能になる。ネットワークは行動経験により得られたデータを用い以下の評価関数を最小化するように重み決定の学習が行なわれる。

$$J = 1/2 \int (\hat{p} - p)^T (\hat{p} - p) \quad (3)$$

学習は、誤差逆伝搬法などを用いて行なわれる。

3.2 プランの生成

プランの目的は、現在地点から目的のセンサーイメージの地点までのパスを生成するモータプランを作成することにある。しかし、このようなパスは無限にあり、問題の設定はこのままでは ill-posed である。そこで、最小距離でといった制約をあらたに設けることにより、解は一意に決まる。問題は以下の評価関数に関する最適化の問題となる。

$$E = \epsilon E_g + \gamma E_c + \mu E_m \quad (4)$$

$$E_g = 1/2 (p^{r*} - p_\tau^r)^T (p^{r*} - p_\tau^r) \quad (5)$$

$$E_c = 1/2 \sum_{n=1}^{\tau} (p_n^l)^2 \quad (6)$$

$$E_m = - \sum_{n=1}^{\tau} \int_0^{x_n} [\phi((x - 0.5)/T) - x] dx \quad (7)$$

ここで、 τ は暫定的な look-ahead のステップ数（つまり、ゴール地点までのステップ数であるが、これは計算の過程で決まってくる）、 p_τ^r はその look-ahead の先端でのセンサーイメージの予測、 p^{r*} は、与えられた目的地点のセンサーイメージである。 E_g は、現時点でのモータプランに基づく、予測センサーイメージとあるべきセンサーイメージのノルムを表す。 E_c は、移動距離に関するコストの予測であり、分岐点間の距離の自乗の総計で表される（ヤマビコは、分岐

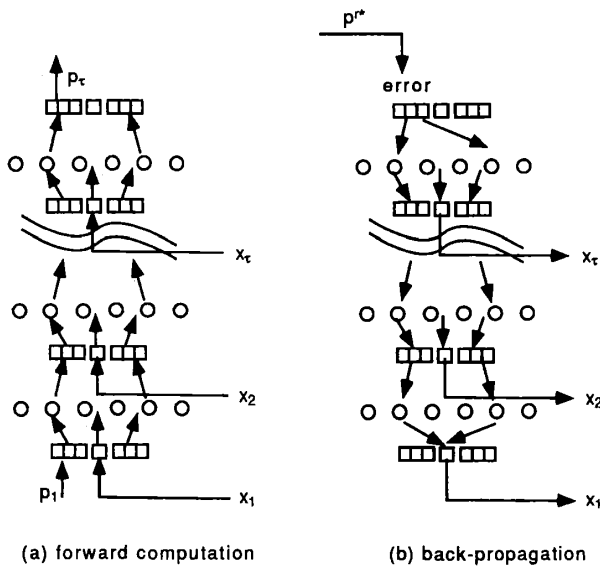


Figure 4: Motor planning in the cascaded RNN. (a) shows forward computation of the distal error by the temporal motor program. (b) shows the back-propagation of the obtained error that updates the motor program.

点間の局所的な移動距離 p_n^l をエンコーダの積分で計算でき、フォワードモデルはセンサー入力の一つとしてそれを予測する)。 E_m は、モータコマンドが分岐情報として 0/1 の値を出すようにするための制約項である。

このような、評価関数を最適化するモータコマンド列 x_n は、図 4 にしめすように、元の RNN を時間方向にカスケードしたものに目的地点のセンサーイメージ p^* を教師信号として逆伝搬をかけることにより繰り返し計算の結果求まる。 x_n の繰り返し計算における更新の方法は以下の式に基づく。

$$\delta x_n = -\epsilon \frac{\delta E_g}{\delta x_n} - \gamma \frac{\delta E_c}{\delta x_n} + \mu(\phi((x_n - 0.5)/T) - x_n) \quad (8)$$

3.3 カオスによる探索

ラチス上での任意のノード間を結ぶ最短距離のパスを決定する問題は多項式オーダーの探索問題であることが、一般に知られている。このことは、式 5 の評価関数は多項式オーダー数の極小解を最適解の他に含むことを示唆し、式 8 の単純な勾配法に基づく解の探索は、簡単に極小解にトラップされることが、予測される。

筆者は、人間が行なうようなトライアンドエラーに基づく deliberative な計算の基盤にはカオス力学系に基づく nontrivial な計算のメカニズムが存在するという立場を取り、本問題にも同じ思考を適用する。具体的には、筆者が提案してきたカオスの最急降下法 (CSD)[1] を、探索メカニズムに組み込み、多様なしかも有効なプランがカオスに駆動され自律的に発生してくるといった仕組みを目指す。CSD は、その特性が正抵抗から負抵抗に徐々に変化する変動型非線形抵抗を最急降下法に組み入れた力学モデルである。その解の探索ダイナミクスは、パラメータの設定によりインターミットントカオスを発現し、ある選択枝をその近傍で細かく探索した後、別の遠くの選択枝にバースト的に移るといった人間の記憶探索過程に良く似た挙動を見せる。(津田、P.Davis 等の Memory Dynamics Model においても、それは顕著である。)

CSD を式 8 に適応し、いかの時間発展式を得る。

$$m\ddot{x}_n + R(\dot{x}_n, \omega t) = -\epsilon \frac{\delta E}{\delta x_n} \quad (9)$$

$$R(\dot{x}, \omega t) = [d_0 \sin(\omega t) + d_1] \dot{x} + d_2 \dot{x}^2 \operatorname{sgn}(\dot{x}) \quad (10)$$

4 実験

適当な、障害物空間を設定し、そこでヤマビコを用いて実験をおこなった。まず、ヤマビコを自律的に障害物回避のみの制御で、実験空間を 30 分ほどランダムに走らせ、193 点の分岐点ごとのセンサー入力及びモータコマンドの時系列データを採集し、RNN に学習させた。図 5 に、光学的に観測した学習過程のトラジェクトリーを示す。

学習が終了したのち、まずロボットを適当に実験空間を周回させ、その間に RNN にセンサーイメージを分岐点毎に入力していき、RNN の隠れユニットを正常な値へと引き込む。つまり、内部のニューラルダイナミクスを環境のダイナミクスによって entrainment することにより、適切なコンテキストを与えさせるわけである。そのあと、図 6 に示すように、目的の移動位置のセンサーイメージを与え、そこへ到達するためのプランニングを開始する。モータプランニングの過程において、カオスダイナミクスに導かれ多様なモータプログラムが時々生

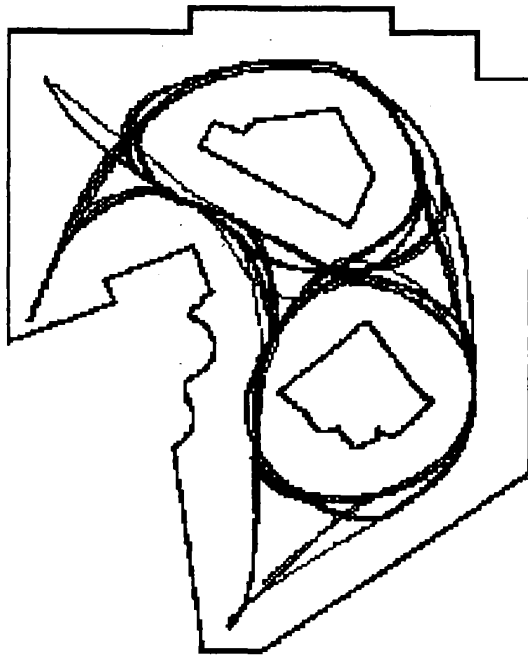


Figure 5: Trace of the latent learning trajectory in the adopted workspace.

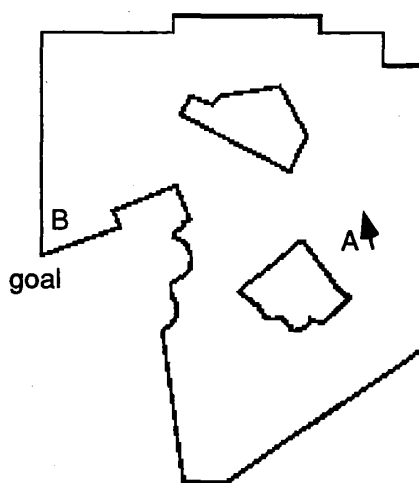


Figure 6: Robot planning for specified goal *B* from current location *A*

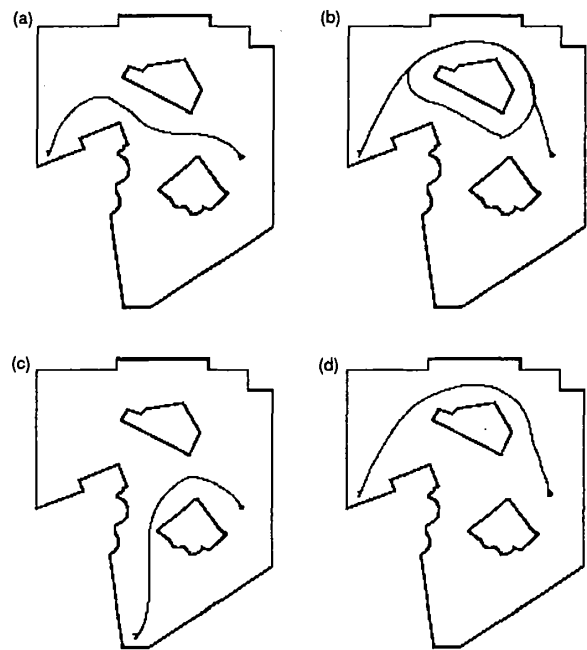


Figure 7: Travels based on generated motor programs. (a) and (d) are almost optimal trajectories, (b) is redundant, and (c) pursues the wrong goal.

成されていく様子が観察された。結局 1 0 1、0 1 0 1、1 1 0 そして 0 0 といった四種類のモータプランが生成された。

これらの得られたモータプランに従って、ヤマビコを実際に走らせてみた。その結果を下図に示す。(a) と (d) はほぼ最適であることが分かる。一方 (b) はゴールに至るものの、冗長な軌道を生じている。(c) では、間違えた地点に到着している。(c) については、間違えた地点のセンサーイメージが目的地点でのそれになっていることが、理由として上げられる。(c) と (b) の評価関数の値は、(a) と (d) のそれより、若干高い。

この実験を繰り返して行なった結果、微小な初期値の違いが全くことなる探索の時間発展を生成することが判明した。しかしながら、多様な解が生成される過程において、最適な解である (a) と (d) は他より、より頻繁に出現することが判明した。このことは、カオスはただランダムではなく、ある程度有効に探索を行なっていることを示している。

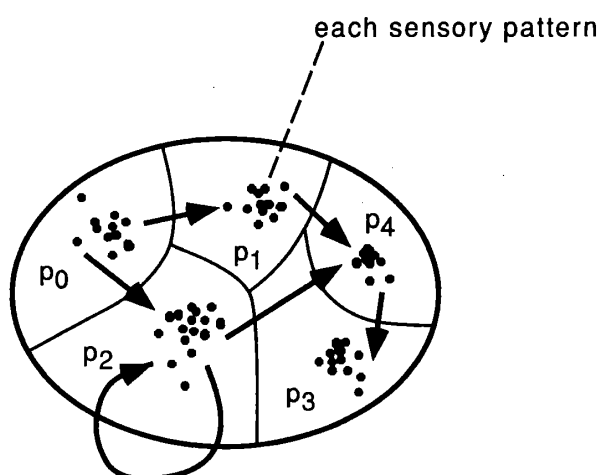


Figure 8: Symbolic sequence of sensory inputs.

5 記号力学系による解釈

ここに示したアプローチを、記号的計算の立場で解釈してみることは、興味深い。実験環境でのナビゲーションにおいて、センサー入力 P は実数ベクトルである。また、先に述べたように、 P は確率的分散を各分岐においてもつ。したがって、ロボットが環境を長い時間動き回ることにより、多数のセンサー入力を得た場合、センサー入力空間において多数の点は環境の分岐構造を反映したような、クラスター構造を作り上げるであろう。各クラスターに記号を与えることにより、センサー入力パターンは量子化される。これより、ロボットは経験を通して、センサーパターンを記号列として受け、そしてその因果関係をフォワードモデルに埋め込んでいくという解釈が成り立つ（図 8 参照）。

ロボットが、プランを立てていく過程を、記号操作の過程として捉えてみよう。まず、ロボットの現在の内部状態“context”を S とする。この時、モータコマンド 0/1 はそれぞれ違った状態 S_0 と S_1 へと内部状態を遷移させる。この時点で、それぞれ P_0 及び P_1 の（記号化された）センサー入力の予測を生成する。これは、以下のような正則文法にて表現できる。

$$S \rightarrow S_0 P_0 (\text{Prob}_0) \mid S_1 P_1 (\text{Prob}_1)$$

ここで、 Prob_0 と Prob_1 は、それぞれ S_0 、 S_1 に遷移する確率であり、それらは移動距離によって決まる（より距離が短い状態への遷移確率はより高い）。

これまでのことにより、フォワードモデルの学習は遷移確率つき正則文法を獲得することと等価であり、一方カオスによるプランニングの過程は、得られた文法を展開して得られるところのマルコフ場において、現状態から、目的状態へと状態遷移していく過程と等価であることが分かる。ここでのカオスの時間発展は文法に基づく形での記号操作の過程と等価である。本論文で特に重要なことは、記号とは用意されたものではなく、物理世界での系の時間発展に対して観測という行為を働きかけた時、自然と発現するものであるという認識であろう。

6 結論

移動ロボットの学習及び行動の例を持って、モデル学習、プランニングといった知的計算の過程を、力学系の体系にて記述する試みを示した。特にプランニングといった、探索的な計算過程は、カオス力学系の時間発展というかたちで記述できることを示した。また、記号とは与えるものではなく、内部的計算のダイナミクスと物理的環境のダイナミクスの密なる干渉の結果、カオス力学系を母体として自己組織化されるものという考えを示した。尚、詳細な報告は [3, 2] 等を参照されたし。

References

- [1] J. Tani. Proposal of Chaotic Steepest Descent Method for Neural Networks and Analysis of Their Dynamics. *IEICE Trans. Fundamentals*, Vol. J74-A, No. 8, pp. 1208–1215, 1991.
- [2] J. Tani. Model-based learning for mobile robot navigation from a dynamical system's perspective. Technical Report SCSL-TR-94-019, Sony Computer Science Lab., 1994.
- [3] J. Tani and N. Fukumura. Learning Goal-directed Sensory-based Navigation of a Mobile Robot. *Neural Networks*, Vol. 7, No. 3,, 1994.